DOT/FAA/ARXX-XX

Office of Aviation Research
Washington, D.C. 20591

Final Report: Visibility in the Aviation Environment

FAA Sponsor Organization: General Aviation, Vertical flight, and Aviation Maintenance Human Factors Program

December 2006

Final Report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AR-xx/xx		

4. Title and Subtitle	5. Report Date
Final Report: Visibility in the Aviation Environment	12/28/2006
	6. Performing Organization Code
7. Author(s) Michael A. Crognale	8. Performing Organization Report No.
9. Performing Organization Name and Address University of Nevada, Reno Reno, NV 89557	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. 03-G-0003
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Aviation Administration Office of Aviation Research Washington, DC 20591	13. Type of Report and Period Covered Final Report 4/3/2003-10/28-2006
	14. Sponsoring Agency Code AFS-800

15. Supplementary Notes

The FAA Technical Center Monitors were William Krebs and Susan Parson

16. Abstract

This report represents studies of visual detection and development of educational materials. The educational materials are designed to help pilots improve their ability to detect other aircraft and to make better informed decisions regarding visibility issues in the cockpit. Results from a number of experiments and detection modeling are presented here. A cockpit aid to detection, a manuscript, and a Powerpoint presentation are among the major final products.

17. Key Words	18. Distribution Statement
Vision, Human Factors, Detection	This document is available to the public through the National Technical Information Service (NTIS), Springfield, Virginia 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	34	

Form DOT F1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGEMENT

We would like to acknowledge the contribution and commitment of the following individuals and their organizations to this effort: Yoko Mizokami, Joe Jaquish, and Kyle McDermot. We would also like to thank our FAA contacts William Krebs (now with the USN), Susan Parson, and Anne Graham for their support and assistance in completing the project.

Table of Contents

Abbreviations

ATC	Air Traffic Control
CFIT	Controlled Flight Into Terrain
FAA	Federal Aviation Administration
FTD	Flight Training Device
GA	General Aviation
IMC	Instrument Meteorological Conditions
PCATD	PC based Aviation Training Device
VMC	Visual Meteorological Conditions

Introduction

General

Visibility issues are a factor in a large number of general aviation accidents each year. Reduced visibility from continued flight into instrument meteorological conditions (IMC) often results in controlled flight into terrain (CFIT), or collision with ground-based obstructions and other aircraft. Poor visibility also is a factor in runway incursions and ground-based accidents. Complex and high contrast backgrounds also contribute to many mid-air collisions by reducing the visibility of other aircraft. Many of these accidents occur in clear skies. Pilots often do not immediately recognize situations that may lead to poor detection and otherwise unsafe visual conditions and therefore fail to take appropriate action. This project has two main goals: 1) trying to better understand visual limitations under conditions of low visibility and decreased detection, and 2) to teach pilots how to better detect other aircraft and to more easily recognize unsafe visual conditions.

Background

Each year there are a large number of accidents in general aviation that result in controlled flight into terrain (CFIT) or collision with other aircraft or land based obstructions such as radio towers (Khatwa& Roelen,1996; O'Hare & Owen, 2002; Volpe, 1994). These accidents occur not only when there is continued visual flight into instrument meteorological conditions (IMC), but often times in conditions of clear weather (reviewed by Kraus, 1995; O'Hare & Owen, 2002). The problem of not being able to visually acquire other aircraft and terrain has its roots in several important issues two of which are considered here.

1) *Learning to see the target-* Visual detection is an active task rather than a passive one. Much of our visual detection is based on "top-down" processing. That is how we see the world depends upon what we have learned to see through experience. Just as when a pilot is learning how to communicate on the radio, knowing what to expect to hear, in large part determines our ability to understand the radio transmissions. Poor audio conditions, external noise, and abundant distractions contribute to the problem of comprehending the transmissions. Efficient search and detection also requires that the observer know what to look for, that is approximately where, when, and how it will appear. The solutions to these tasks are easily calculated from known relationships. Training is required however for pilots to perform quickly and automatically. The present study has produced a cockpit aid to traffic detection (described in detail below) that should help many pilots improve traffic detection. We also describe an experimental program that helps train pilots to recognize traffic target altitudes, direction of travel, and distance.

2)Learning to judge the visual environment- There are three components to this issue a) the background, b) intervening atmosphere and c) lighting especially "flat-light".

The background against which targets must be detected varies from low contrast, uniform (e.g. clear blue sky) to complex and high contrast (e.g. cityscapes and mottled mountainous terrain). In general, detection is inversely related to scene complexity. In other words, the more complex and higher contrast the background, the harder it is to detect a target on it. This is a phenomenon known as masking in the vision science world.

There are various ways to characterize scene complexity and analyze the statistics of visual scenes. One of these techniques involves Fourier analysis which breaks down an image into component spatial frequencies and a subsequent plotting of the resultant spatial frequency amplitude and phase spectra. Although it has been argued that most natural images show spatial frequency spectra that fall off in amplitude as 1/f, there is ample evidence that the spectra of many scenes differ from 1/f significantly (e.g. Field & Brady, 1997).

An alternative to Fourier analysis involves the application of sparse coding algorithms (Simoncelli & Olshausen, 2001) to images from the aviation environment. This algorithm produces basis functions which are believed to be generated in a similar manner to the receptive fields of visual cortical neurons, that is, by learning from the statistics of the environment. Such an application provides insight as to the limits of applying our land based visual system to the demands of the aerial environment. In the present work we present data on the statistics of images from the aviation environment and compare those with statistics from images of terrestrial scenes.

Detection models and performance

The sparse coding algorithms discussed above suggest a new model of detection based upon the differences between the basis function weightings for a target and those of the local background. Specifically, it would be predicted that the more different the weightings of basis functions for targets are from those of the local background, the easier the target should be to detect. There are currently numerous models that take into account local statistics such as contrast such as that proposed by Ahumada (1996). In the present study we have collected detection data using backgrounds of images from the aviation environment and compared model predictions from several versions of our new model descried above and Ahumada's model with the detection data.

Improvements in detection with the aid of lighting

External aircraft lighting, in particular strobe lights have been shown to greatly improve detection. When lights in different areas are flashed out of sequence a percept of motion can arise. The improvement of detection in this case arises by virtue of stimulation of a visual pathway specifically tuned to transient changes such as flashes or motion. The strength of this motion percept depends greatly upon parameters such as the timing of flashes and the distance by which they are separated. In this study we present data on how these parameters contribute to the detection of targets on backgrounds of various complexity.

Pilot education

We have developed a Powerpoint presentation with the goal of pilot education regarding visibility and strategies for improving detection. It is our goal to present this talk at numerous aviation gatherings to reach as many pilots as possible. In addition the presentation is annotated such that it can be available for flight schools and other interested parties to make the presentation. This presentation is described in more detail below and should become available through the FAA.

Results and Products

Simulator

We have completed construction of a flight simulator (PCATD) with extended visual display (see figure 1). The flight simulator is approved for instrument instruction and basic flight instruction as outlined by the Federal Aviation Regulations. This simulator is currently used in experiments of visual detection and provides nearly 180 degrees of visual field of view. We plan to fully utilize the simulator's capabilities to answer questions regarding the influence of factors such as fatigue and inclement weather on detection and pilot performance.

Figure 1. Flight simulator developed for visual detection and human factors experiments.

Figure 2 shows four views from one of the panels of the PCATD with a target aircraft at different altitudes and distances. The subject's task is to press a button on the yoke when the target is detected.

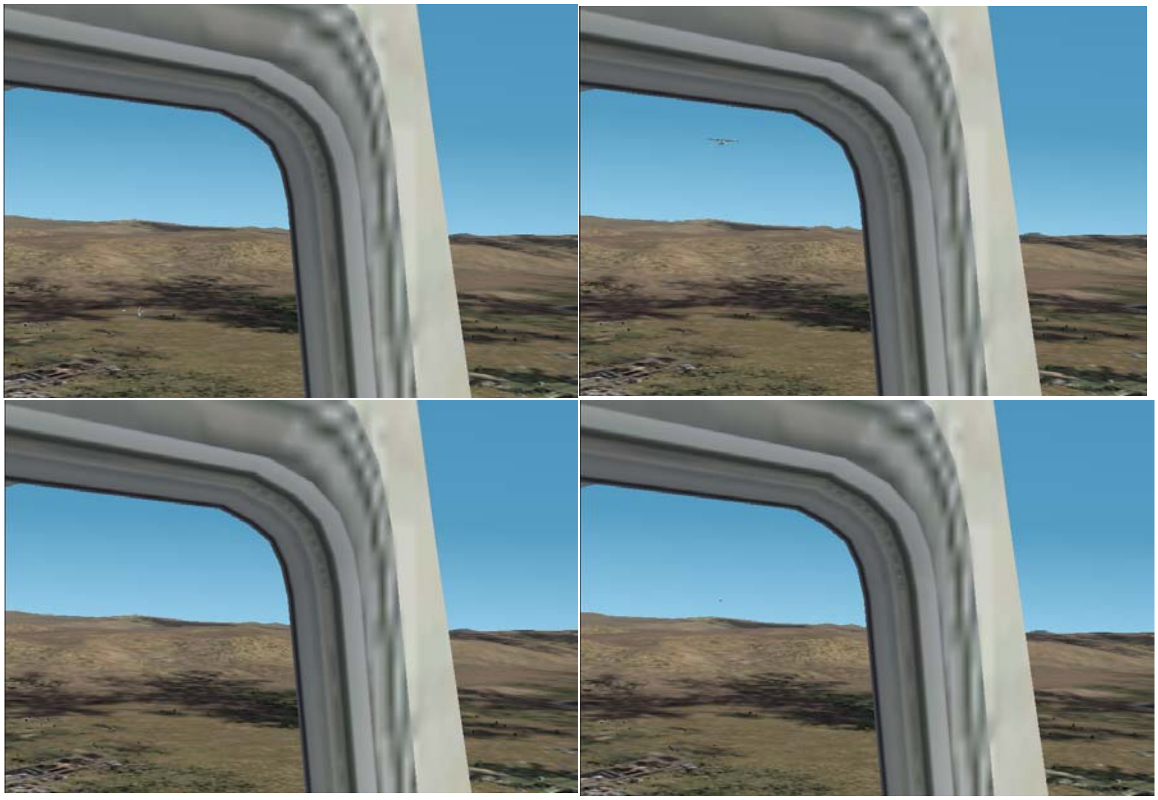


Figure 2. Images showing one panel of the PCATD during a detection experiment. Images on the top show a Cessna at a far distance while the images in the bottom show closer aircraft. The target aircraft in the left hand panels are below the horizon while those on the right are above the horizon.

Sparse coding of aerial images.

Intuitively, the aerial visual environment may appear to be quite different from the terrestrial visual environment. When looking out of an aircraft for example, there is a wide view of the sky and geographic features such as mountains. Objects on the ground appear tiny. However, there is a wide variety of possible image characteristics in both aerial and terrestrial environments and it is unknown how the characteristics of images from the aerial environment differ quantitatively from those of images from the terrestrial environment. A quantitative knowledge of such differences may be useful for understanding and modeling detection and visual performance for operations in novel visual environments.

Here we use a sparse coding model (Olshausen & Field 1996, 1997) to characterizing both terrestrial and aerial images. This approach has been shown to generate responses that are similar to those observed in some cells of the primary visual cortex, providing a novel method to characterize natural images in relation to the activities of the visual system. In particular, the spatial properties of receptive fields in primary visual cortex have been characterized as localized, oriented, and bandpass, comparable with basis functions derived from natural images (Olshausen & Field 1996, 1997). The properties of these fields may arise from the strategy of producing a sparse distribution of neural activity in response to these images. This suggests that sparse coding could characterize the properties of images in relation to responses in the cortex.

Although sparse coding models derived from terrestrial-based images predict the spatial characteristics of receptive fields well, these analyses have been based on images sampled from the terrestrial environment and have not yet been applied to the aerial environment. It may be of interest to determine whether or not application of the same sparse code algorithm to images from the aerial environment would result in derivation of similar spatial characteristics of receptive fields (basis functions) as those derived from terrestrial images. Consideration of the complementary issue, that is, whether or not basis functions derived from terrestrial images can be used to adequately describe images from the aerial environment, should also provide insight into adaptation of visual systems to novels environments. Differences in the relative amounts of each basis function required to encode the information for both aerial and terrestrial images may provide insight as to the adaptive capabilities of visual systems that utilize sparse-coding-like "learning" mechanisms.

Here we investigate whether characteristics of images in the aerial environment differ from those of terrestrial-based images and how such differences might affect visibility. In the first analysis, we compared the characteristics of images of aerial and terrestrial environments using the sparse coding technique. In order to further characterize the adaptive response of the cortex to the aerial environment we applied basis functions learned from terrestrial images to both terrestrial and aerial-based images, and compared the weightings (coefficients) of the basis functions for these two classes of images.

Method

We applied a program which incorporates the sparse coding algorithm described by Olshausen and Field (1996) to natural images of both terrestrial and aerial environment. The aerial images were obtained from the cockpit of a Cessna 206 and were taken over a period of 18 months and in approximately 12 states in the U.S. including Alaska. Images were taken with digital cameras. Terrestrial images were taken over a period of 24 months mainly in Reno and the surroundings but also in Japan and India.

Image model

The model starts with the basic assumption that an image, $I(\vec{x})$, can be represented in terms of a linear superposition of basis functions $\phi_i(\vec{x})$, with amplitudes a_i

$$I(\vec{x}) = \sum_i a_i \phi_i(\vec{x}) \qquad (1)$$

The image code is determined by the choice of basis functions ϕ_i. The coefficients, a_i, are dynamic variables that change from one image to the next. They are computed for each image to satisfy the above equality, and these quantities constitute the output of the code. The goal of efficient coding is to find a set of ϕ_i that forms a complete code and results in the coefficient values being as statistically independent as possible over an ensemble of natural images.

Olshausen and Field suggest that natural images have 'sparse structure' that is any given images can be represented in terms of a small number of descriptors out of a large set. A specific form of low-entropy code is sought in which the probability distribution of each coefficient's activity is unimodal and peaks around zero.

The search for a sparse code is formulated as an optimization problem by constructing the following cost function to be minimized:

$$
\begin{aligned}
E &= -[\text{preserve information}] - \lambda[\text{sparseness of } a_i] \\
&= \sum_{\vec{x}} \left[I(\vec{x}) - \sum_i a_i \phi_i(\vec{x}) \right]^2 + \lambda \sum_i S(a_i)
\end{aligned} \tag{2}
$$

Where λ is a positive constant that determines the importance of the second term relative to the first. The first term measures how well the code describes the image. $S(a_i)$ is a nonlinear function (cost function). Learning is accomplished by minimizing the total cost function, *E*. For each image presentation, *E* is minimized with respect to the a_i. The ϕ then evolve by a gradient descent on E averaged over image presentations. The a are determined from the equilibrium solution to the differential equation:

$$
a_i = \sum_{\vec{x}} \phi_i(\vec{x}) r(\vec{x}) - \lambda S'(a_i), \tag{3}
$$

Where $r(\vec{x})$ is the residual image

$$
r(\vec{x}) = I(\vec{x}) - \sum_i a_i \phi_i(\vec{x}) \tag{4}
$$

The first term of equation (3) takes a spatially weighted sum of the current residual image using the function ϕ_i as the weights. The second term applies a non-linear self-inhibition on a_i, according to the derivative of *S*, that differently pushes activity towards zero.
The learning rule for updating ϕ_i is then:

$$
\Delta \phi_i(\vec{x}) = \eta \langle a_i r(\vec{x}) \rangle \tag{5}
$$

Where η is the learning rate.

This algorithm seeks a set of basis functions ϕ_i for which a_i can tolerate 'sparsification' with minimum reconstruction error to find a set of ϕ_i that can best account for the structure in the images in terms of a linear superposition of sparse statistically independent events. The basis set is overcomplete, meaning that there are more basis functions than effective dimensions in images. Overcompleteness in representation is important because it allows for the multidimensional space of position, orientation and spatial-frequency to be tiled smoothly without artifacts. More generally though, it allows for a greater degree of flexibility in the representation, as there is no reason to believe a priori that the number of causes for images is less than or equal to the number of pixels.

Simulation methods

In one image-set, the data for the training process were taken from thirty 512 x 512 pixel images. Training data were obtained by extracting 12 x 12 pixel image patches at random from images that were preprocessed by filtering with the zero-phase whitening/low-pass filter $R(f) = fe^{-(f/f_0)^4}$, f_0 = 205 cycles/picture. Whitening counteracts the fact that the mean-square error preferentially weights low frequencies. Basically corrects for the vast differences in variance across spatial-frequencies due to the $1/f^2$ power spectrum of natural images. This process simulates the filtering done by the retina and LGN (Atick & Redlich, 1992), and also speeds the learning process.

a_i is computed by first initializing to

$$a_i^0 = \sum_{\vec{x}} \phi_i(\vec{x}) I(\vec{x}) \quad (6)$$

and then iterating equation (3) using the conjugate gradient method, halting after 10 iterations, or when the change in E is less than1%.

The ϕ_i were initialized to random values and were updated every 100 image presentations (patches). The vector length (gain) of each basis function, ϕ_i, was adapted over time so as to maintain equal variance on each coefficient. 50,000 updates were taken for a stable solution. The parameter λ was set so that $\lambda/\sigma = 0.14$ with σ^2 set to the variance of the images. The form of the sparseness cost function was $S(x) = \log(1+x^2)$.

The sparse coding process was run for 12 image-sets for the terrestrial and aerial environments, respectively. Fig. 1 (a) shows examples of image-sets from each environment: (a-1) to (a-3), terrestrial; (a-4) to (a-6), aerial. Each image-set consists of thirty images. We prepared image-sets including different sceneries such as forest, mountain, desert, city, and the university campus for the terrestrial environment. Aerial images included both forward and down views of various terrain, geographical features, and structures.

Results

An example of a derived basis function set is shown in Fig.3 Coefficients for each of these functions can describe any arbitrary image from the "learning" set of images.

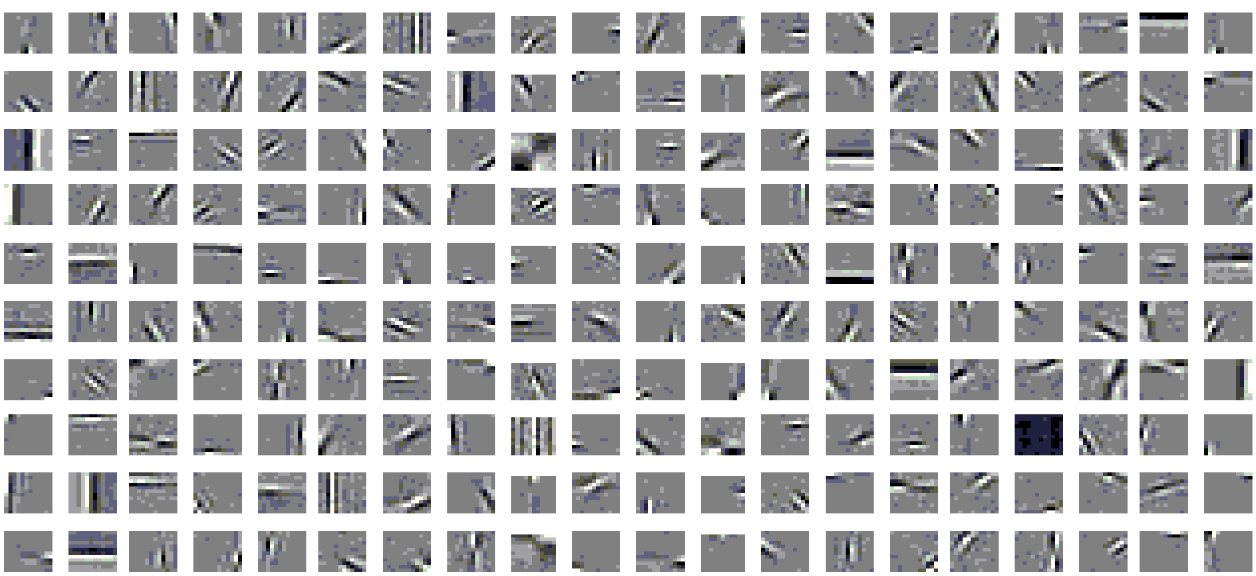


Figure 3. Set of basis functions "learned" from a set of aviation images (for example those from figure 4 (a4-a6). The relative weights (coefficients) of each of the basis functions can be used to describe any particular scene.

Fig. 4 (b-1) to (b-6) show the outputs of basis functions from image-sets (a-1) to (a-6), respectively. The basis functions from terrestrial images show clear Gabor function shapes (well localized, oriented, and bandpass) which are consistent with Olshausen et al.'s results. Overall, the basis functions learned from aerial images are noisier than those from terrestrial images suggesting that the characteristics of aerial images may differ from those of terrestrial scenes. Additionally, some aerial image sets (e.g. Fig.4 b-4) did not converge on meaningful basis functions most probably due to a paucity of high spatial frequency components (e.g. including only hazy images and clouds).

To analyze the characteristics of the derived basis functions, we computed the power spectra of the functions to obtain their peak spatial frequency and orientation. For the position parameter we computed the Hilbert transform and took the peak of the modulus of the quadrature pair. As shown Fig. 5, for example, the basis functions (b) learned from a terrestrial image-set (a) have different distributions of orientation, position, and spatial frequency. The number of basis functions with a particular orientation and spatial frequency were calculated and are shown in Fig. 6 (a). The orientation of each basis function was rounded off to the nearest 10 degree step and binned into 10-degree-steps. The same procedure was applied to spatial frequency using steps of 1 cycle/image (cycle/12 pixel). The derived basis functions are well distributed especially for vertical (0 deg.) and horizontal orientations (90 or -90 deg.). In the case of spatial frequency, the peak of the distribution is at 4 cycle/image.

Coefficients for 100 image patches in each basis functions were obtained as the output of the sparse coding algorithm. One set of coefficients is shown in Fig. 6 (b). This figure illustrates the sparseness in that not all basis functions are significantly active. To obtain the overall trend of excitation for the image-set, we took a root mean square (rms) of 100 coefficients in each basis function as shown in Fig. 6 (c). The rms of coefficients for 200 bases are re-plotted against orientation and frequency as shown in Fig. 3 (d). Each point corresponds to the rms of each basis function. Their distributions show the same trend as the distribution of the number of basis functions (a). This means that the basis functions are well distributed but more active in the vertical and horizontal orientations. In the case of spatial frequency, the points that have high coefficient values are rather scattered. Since many of the points in (d) overlap each other, we took a sum of the rms in each orientation (10 deg. steps) and spatial frequency (1 cycle/image-patch steps), respectively, as shown in Fig.6 (e). These data provide an indication of the overall activity of the basis functions taking into account both the number and the value of the coefficients. We calculated the output from all the image-sets in the same way and compared the characteristics of the images. It should be noted, though, that the basis functions from 5 aerial image-sets were eliminated because they did not converge to Gabor shapes and could not give meaningful power spectra.

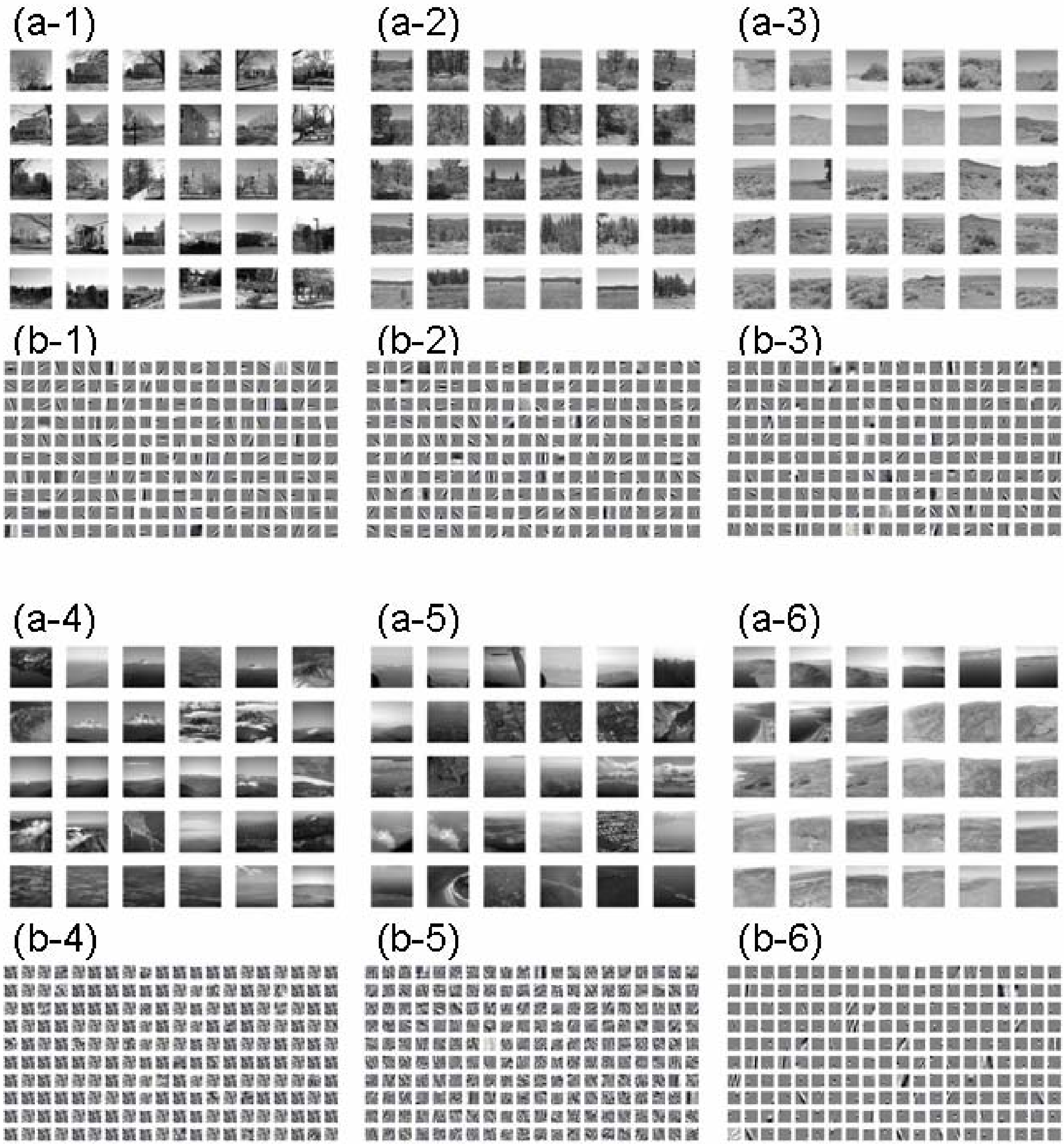


Figure 4 Examples of image-sets and basis functions. (a-1) ~ (a-3), image-sets from terrestrial environment; (a-4) ~ (a-6), image-sets from aerial environment. (b-1) ~ (b-6), 200 basis functions delivered from each image-set.

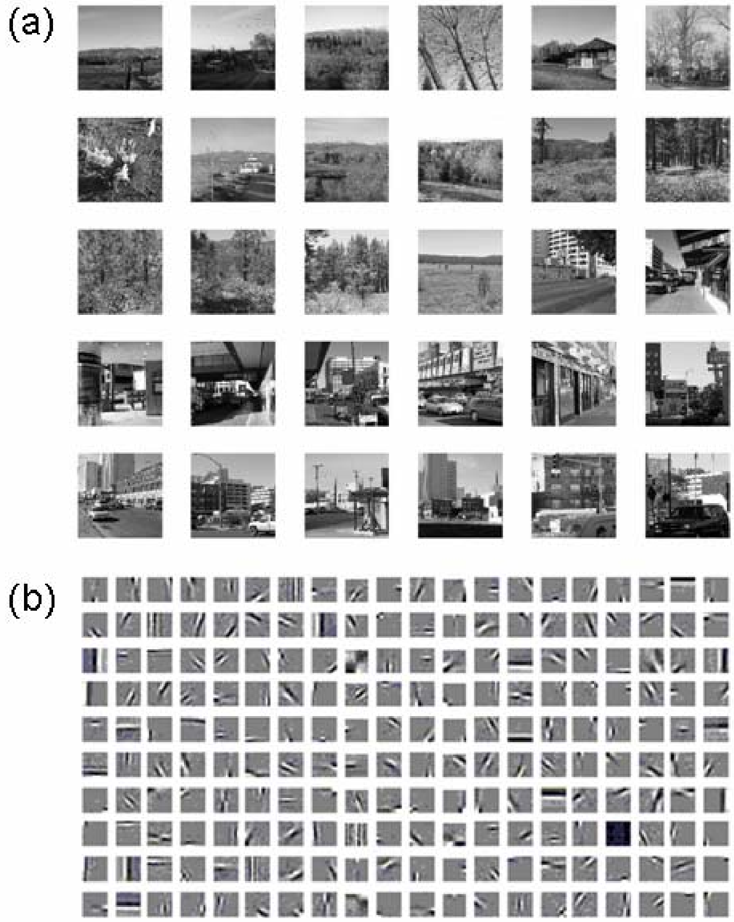


Figure 5 (a) An image-set from a terrestrial environment (b) Basis functions from image-set (a). These functions were used for the procedure in section 4.

The results of the analysis for all the image-sets (except for the 5 aerial image-sets that failed to converge) are shown in Fig. 7. Fig 7 (a) and (b) show the results from12 terrestrial image-sets and those from 7 aerial image-sets, respectively. For all the terrestrial image-sets, the distributions of orientation have the same shape and clearly show more vertical and horizontal orientations (a). The distributions for the aerial images have no clear peak in any specific orientation (b). The distributions of spatial frequency also show clearer peaks for the terrestrial images than for the aerial images. Those for the aerial images are rather flat. Fig. 7 (c) and (d) are the averages plotted with standard deviations. Overall, the standard deviations for the distributions of the terrestrial images are smaller than those of the aerial images. This implies that there is more variation from one image set to another in the aerial environment and the aerial image sets may be harder to characterize. The sums of coefficients are generally smaller for the aerial image sets implying that those particular basis functions are less active. The distributions of position information did not show systematic differences between the terrestrial and aerial environments and we chose not to present those data here.

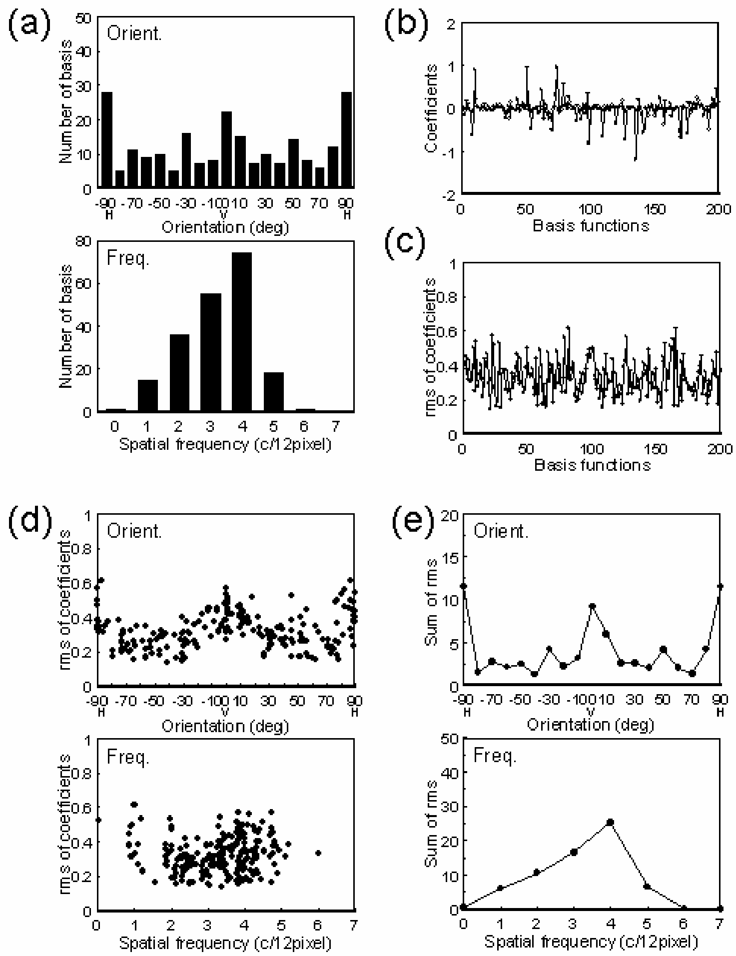


Figure 6 (a) The number of basis functions against orientation (upper graph) and spatial frequency (lower graph). H and V in upper graph stand for 'Horizontal' and 'Vertical', respectively. (b) An example of coefficients. The coefficients of each basis functions for a patch from image-set (a). (c) Root mean square (rms) of coefficients from 100 patches. (d) Rms of coefficients against orientation (upper) and spatial frequency (lower). H and V in upper graph stand for 'Horizontal' and 'Vertical', respectively. (e) Summed rms of coefficients against orientation (upper) and spatial frequency (lower).

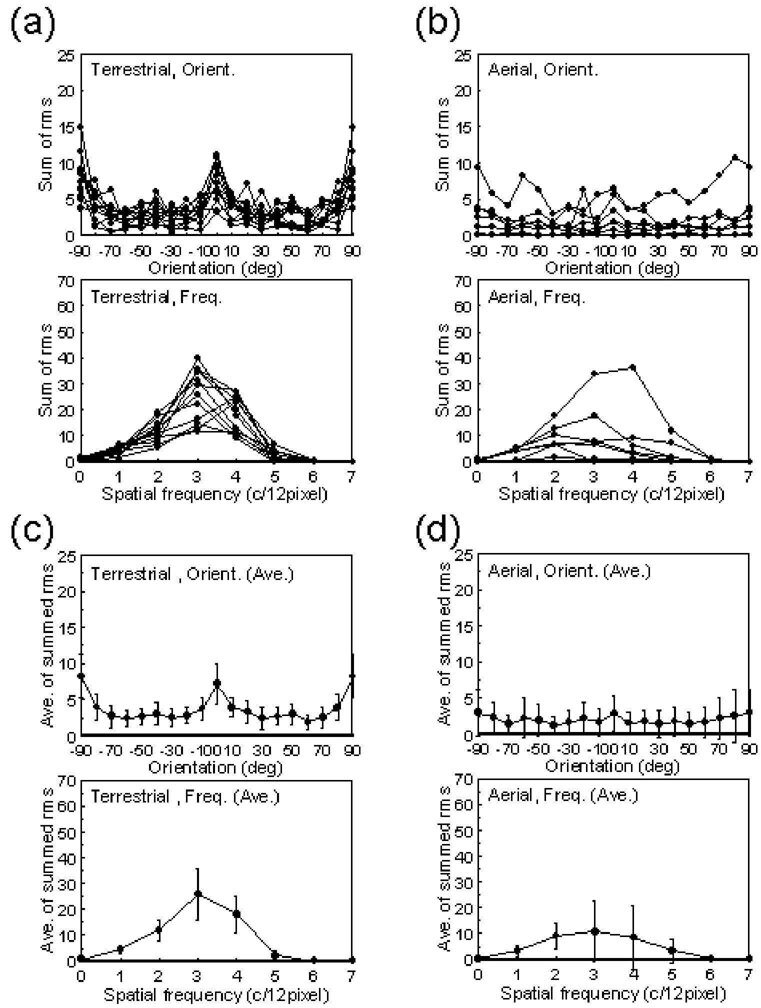


Figure 7 Characteristics of response against orientation (upper) and spatial frequency (lower). (a) Results from 12 image-sets from terrestrial environment. Each series indicate summed rms of coefficients from each image-set. (b) Results from 7 image-sets from aerial environment. Each series indicate summed rms of coefficients from each image-set. (c) Average of the summed rms of 12 image-sets from terrestrial environment. Error bars indicate standard deviations. (d) Average of the summed rms of 7 image-sets from aerial environment. Error bars indicate standard deviations.

Terrestrial basis functions in the aerial environment.

In order to further characterize the simulated response of the cortex to the aerial environment, we applied basis functions learned from terrestrial images to both terrestrial and aerial-based images, and compared the coefficients (weightings) of the basis functions for these two classes. We applied the basis functions derived earlier and shown in Fig. 5 to all other image-sets from the terrestrial and aerial environments. This was done in order to simulate how receptive fields in our visual cortex (derived from learning in the terrestrial environment) might be activated in the aerial environment.

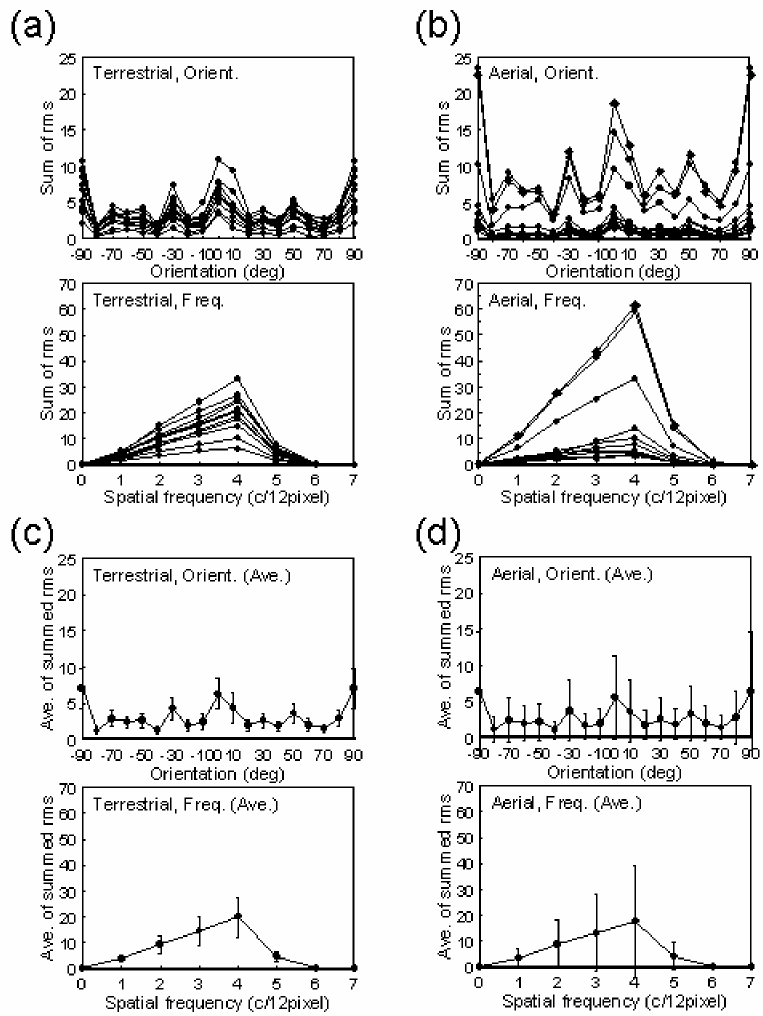


Figure 8 Characteristics of response when same basis functions were applied.
(a) Results from 11 image-sets from terrestrial environment. (b) Results from 12 image-sets from aerial environment. (c) Average of the summed rms of 11 image-sets from terrestrial environment. Error bars indicate standard deviations. (d) Average of the summed rms of 12 image-sets from aerial environment.

We used the same algorithm as described above except that the previously derived basis functions were utilized throughout all the iterations. Thus the modified algorithm sought only the optimal coefficients for those basis functions. It should be noted that to avoid circularity in this analysis, we did not apply the basis functions to the original image set that was used to derive those functions.

The results for 11 image-sets from the terrestrial environment and 12 image-sets from the aerial environment are shown in Fig. 8. We did not included the result of the image-set which was used to derive these basis function because the purpose of this analysis was to examine how these basis functions (receptive fields) learned from one type of environment respond when they are stimulated by new environments. The activity (sum of coefficients) for all the terrestrial image-sets of show consistent trends (a). Those of aerial images differ depending on the particular image-set (b). In most aerial image-sets, the activity remains low. Greater activity is evidenced in three image-sets. It is of interest that these particular image-sets include a number of down-looking views of cities, fields, and mountains (Fig. 4 (a-4) for example) while the other image-sets showing low activity also included similar views. Further analysis is therefore needed to explain the differences of activity. Fig. 8 (c) and (d) show the average of the summed coefficients for images from the terrestrial and aerial environments, respectively. The averages of each environment are similar but the standard deviations are large for the coefficients from the aerial environment. This illustrates larger variability or poorer fit of terrestrial functions to aerial images. The results from analysis of aerial images suggest that our visual system, which is adapted to the terrestrial environment, may not be optimized for the aerial environment.

Image-sets used here were arbitrarily selected from various sceneries; natural environments such as mountains, forests, desert or city, and the university campus for the terrestrial environment, and mountains, city, cloudy sky, etc for the aerial environment. Even though there is a large variety in the content and composition of images in both environments, the resultant basis functions from the terrestrial images showed similar characteristics to each other but those from the aerial environment had large differences. This suggest that for sparse coding type "learning" in cortical receptive fields, the two environments may not be equivalent and vision adapted to terrestrial environments may not be optimized for the aerial environment. Whether or not this lack of optimization has significant or measurable visual consequences needs to be determined. In addition, since the aerial environment differs enough from the terrestrial environment to produce different basis functions and in some case does not present enough structure to "learn" sufficiently to converge on an optimal basis set then it is likely that cortical fields learned in such an environment would perform poorly for terrestrial image processing. Characterization of cortical receptive fields that have developed in such altered and/or impoverished environments may reveal such deficiencies.

Sparse coding models simulate many of the characteristics of cells in the visual cortex successfully, when they utilize natural images from the terrestrial environment. It may be that sparse coding does not work with aerial images in the same manner as with terrestrial images because our brains have been developed and adapted within a terrestrial environment. If human developmental visual mechanisms act like sparse coding processes, the "failure" of coding for a new, unfamiliar environment may be predictable. However it should also be noted that with regard to causation, the failure of some aerial image sets to produce reliable and consistent basis functions must rest with the nature and quantity of the information inherent in these images; the learning is worse in the aerial environment than the terrestrial environment because there is less information (structure).

Our results suggest that aerial images can be characterized by sparse coding model. However, basis functions derived from sparse coding for aerial and terrestrial environments have different characteristics. Basis functions from terrestrial images are consistent, while those from aerial images are not. If cortical receptive fields in humans develop in a manner similar to sparse coding then the aerial environment may be relatively unusual or novel for our visual system, which is adapted to the terrestrial environment. Future work should address whether or not this mismatch results in measurable visual deficits and if rearing in altered environments produces altered receptive fields consistent with those predicted from sparse coding algorithms.

Detection models and performance

As described above we developed several versions of a model of detection based upon the sparse coding algorithm described above. We tested these models as well as a contrast masking model of detection proposed by Ahumada and Beard (1997) against detection data collected in our lab. First we will describe the detection models and then the experimental procedures.

Masking model (from Ahumada 1996)
(Single filter model with masking by a non-homogeneous background)

The following steps to create "visible contrast" images were applied to the images with and without a target.
The input to the model consists of two images. The output is a perceptual distance d', representing the number of just-noticeable-differences between the images. Each of the following steps is applied to both images.

- **Blur**. The image I is convolved with a low pass Gaussian filter F_B (= $e^{-(f/fc)^{\wedge}2}$),
  $B[x,y] = I[x,y] * F_B[x,y]$.
- **Local luminance**. The blurred image B is convolved with a low pass Gaussian filter F_L,
  $L[x,y] = B[x,y] * F_L[x,y]$.
- **Local contrast**. The contrast image is computed from the local luminance,
  $C[x,y] = B[x,y] / L[x,y] - 1$.
- **Local contrast energy**. Squared contrast image values are convolved with a Gaussian low pass filter F_E,
  $E[x,y] = C[x,y]^2 * F_E[x,y]$.
- **Local contrast gain adjustment**. The masked visible contrast image is computed using a divisive inhibition formula,
  $V[x,y] = C[x,y] / (1+g_E\ E[x,y])^{0.5}$.
- **Summation of image differences**. The distance between the masked visibility images is based on a Minkowski metric with an exponent of 4, corresponding to probability summation over space,
  $d' = g_C\ (\Sigma_{x,\,y}\ (V_1[x,y] - V_2[x,y])^4)^{0.25}$.

Some examples of the filtering procedure and resultant predictions are shown in figure 9. The last image represents the probability of detection as a function of lightness. In this case the brighter the region is, the more difficult target detection would be on that region.

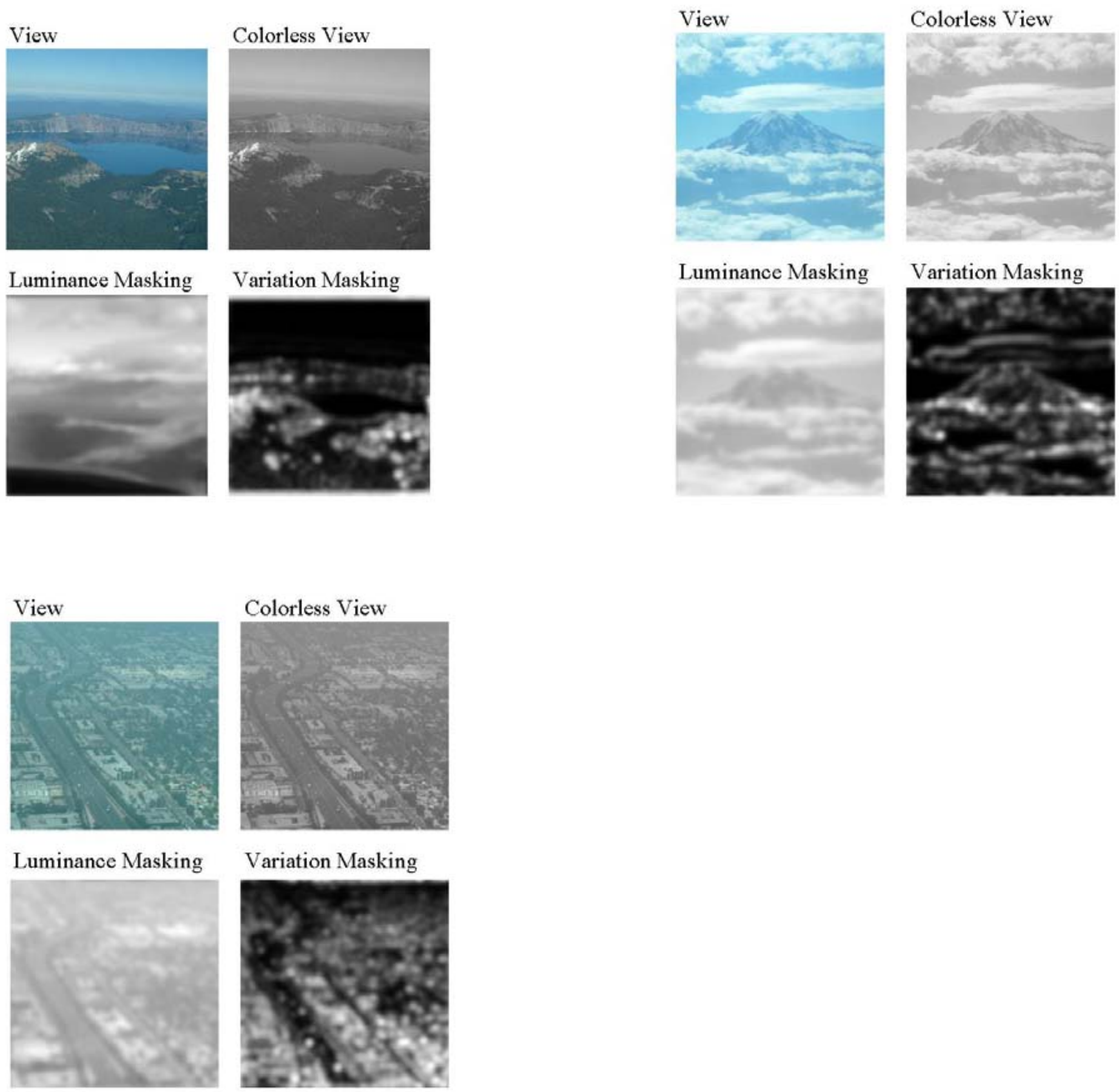


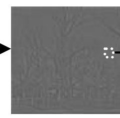
Figure 9. Examples of images filtered using the detection model of Ahumada (1996). The last image represents the probability of detection as the inverse of brightness. See text for model details.

Model from Sparse coding model

1. The basis functions were first calculated from terrestrial images. The sparse coding model starts with the basic assumption that an image, $I(\vec{x})$, can be represented in terms of a linear superposition of basis functions $\phi_i(\vec{x})$, with amplitudes a_i

$$I(\vec{x}) = \sum_i a_i \phi_i(\vec{x}) \qquad (1)$$

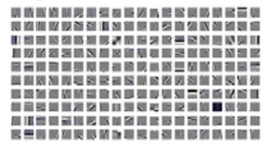
Image set (30 512x512 images)



12X12 patches

$$I(\vec{x}) = \sum_i a_i \phi_i(\vec{x})$$

Whitening → Training → Output



Coefficients

$$a_i = \sum_{\vec{x}} \phi_i(\vec{x}) r(\vec{x}) - \lambda S'(a_i),$$

$$r(\vec{x}) = I(\vec{x}) - \sum_i a_i \phi_i(\vec{x})$$

200 basis functions

ϕ

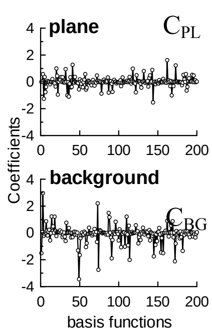
Updating basis function

$$\Delta\phi_i(x_m, y_m) = \eta \left\langle a_i \left[I(x_m, y_n) - \hat{I}(x_m, y_n) \right] \right\rangle$$

2. Coefficients of plane and background

The basis functions obtained above applied to plane target and background in all positions of a background image. The a are determined from the equilibrium solution to the differential equation (same as above except for using the same basis functions).

Plane-image
(12 x 24 pixel)



12X12 patch

BG-image
(1024 x 1280)

3 prediction models

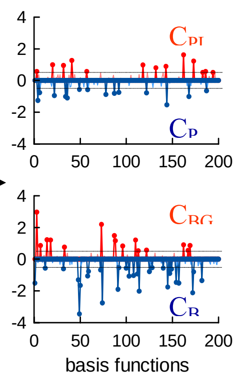
1. Sum $\boxed{\sum(C_{PL}^2 - C_{BG}^2)}$

   Coefficients are summed and compared.

2. ON(OFF)-sum $\boxed{\sum(C_{PL+}^2 - C_{BG+}^2 + C_{PL-}^2 - C_{BG-}^2) / 2}$

3. ON(OFF)-diff $\boxed{\sum(C_{PL+} - C_{BG+})^2 + \sum(C_{PL-} - C_{BG-})^2 \} / 2}$

   Coefficients are separated into Positive/ Negative, and thresholded by SD of $(C_{pl}+C_{BG})$, then compared.



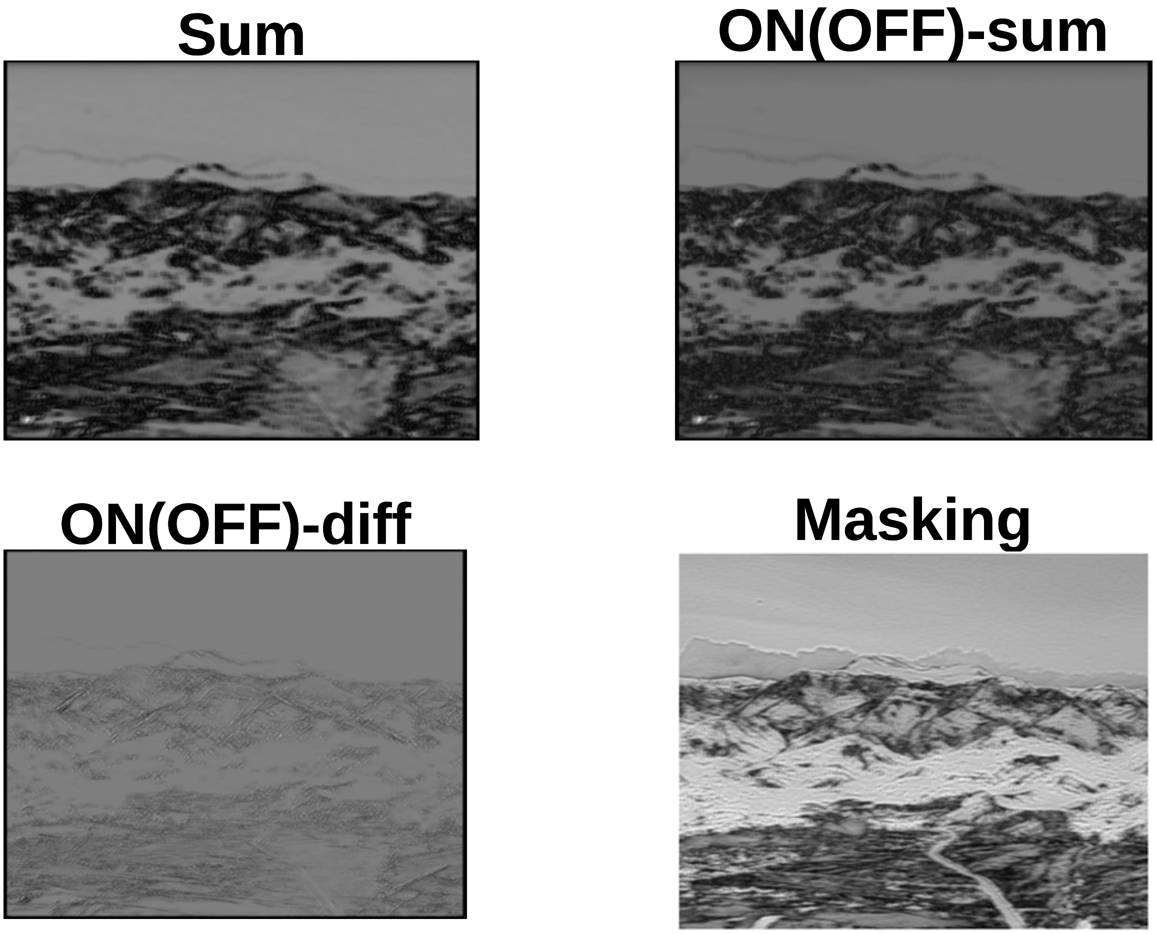


Figure 10. Detection predictions for the 4 models that were tested. In this rendition darker regions correspond to more difficult detection.

Methods.

% subjects were tested. An airplane-shaped target (0.5 deg) was shown on gray images (the aviation environment) on CRT monitor randomly in one of 4 quadrants (see figure 11). Subjects judged in which quadrant the target appeared. Detection and reaction time (RT) were measured using either 180 random positions in 5 images and also using 19 fixed positions: each (at red dots shown on the images in figure 12). The fixed regions were chosen to sample high, and low detection regions and also levels at which the models made maximally different predictions.

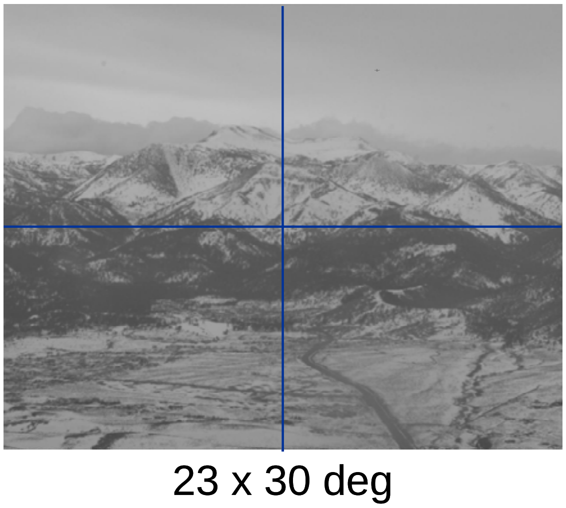


Figure 11. Example of image used in detection task.

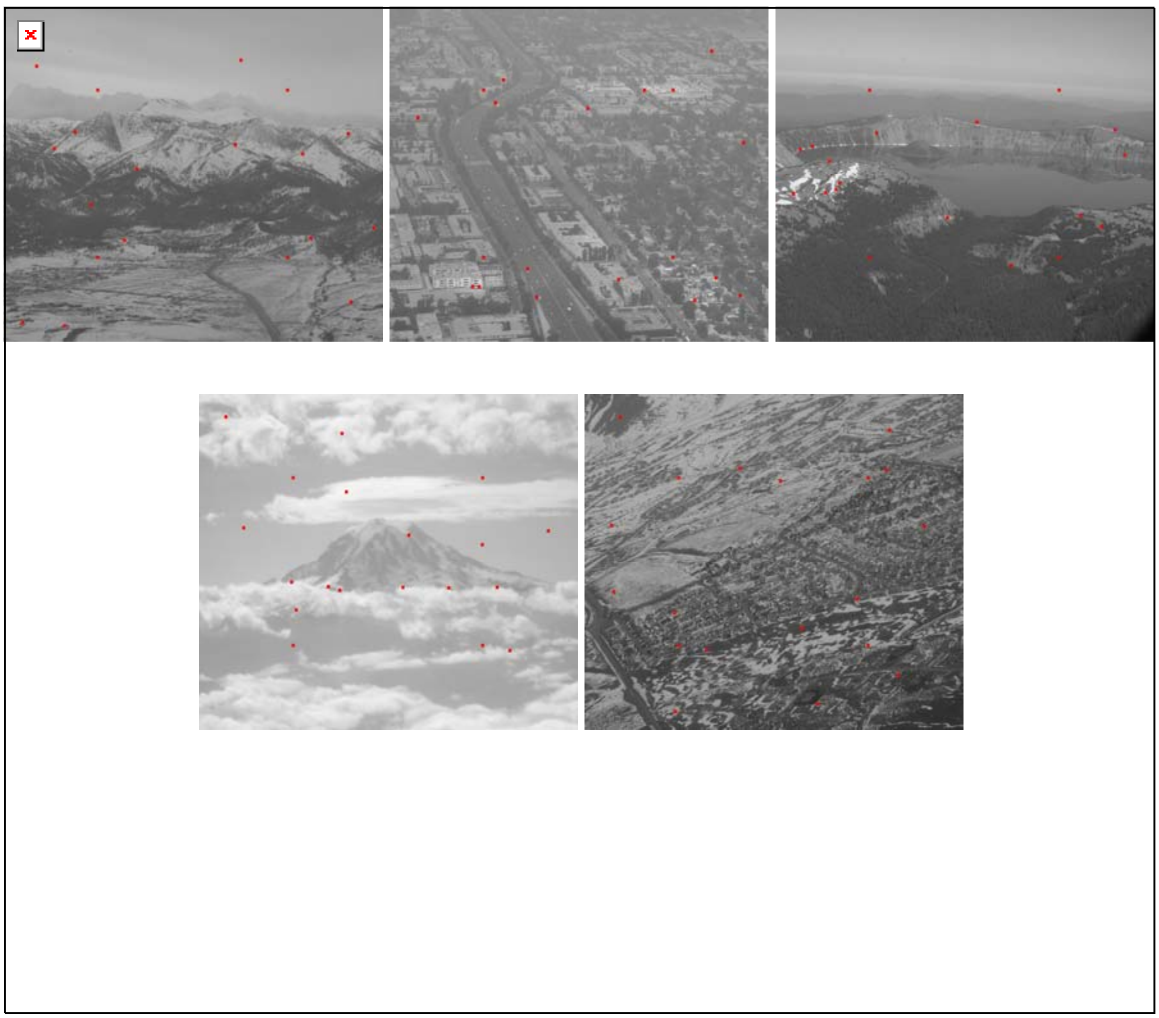


Figure 12. Images and fixed locations chosen for detection tasks

Results

The results from the random positions are shown in figure 13 for 2 subjects. Performance correlated well with predictions for all 5 subjects although there were also examples of large discrepancies between each of the model predictions and the data.

The results from the fixed position tests are shown in figure 14 and 15. The detection and reaction time results are shown in figure 14 while the results of a comparison of the models is shown in figure 15. In this figure we plot the sum of the squared error of the model fits for 5 subjects. We found that although the image analysis based on sparse coding was quite useful for quantifying the image characteristics, the models developed using the algorithm did not provide significant advantage over the mathematically simpler Ahumada and Beard model. The frequent departures from the model predictions in general suggest that other variables not well described by the models need to be accounted for in an improved model of detection.

In the future we plan to test the predictions of the models against behavioral detection results obtained in a more realistic aviation setting include distractions and flying tasks provided by the flight simulator.

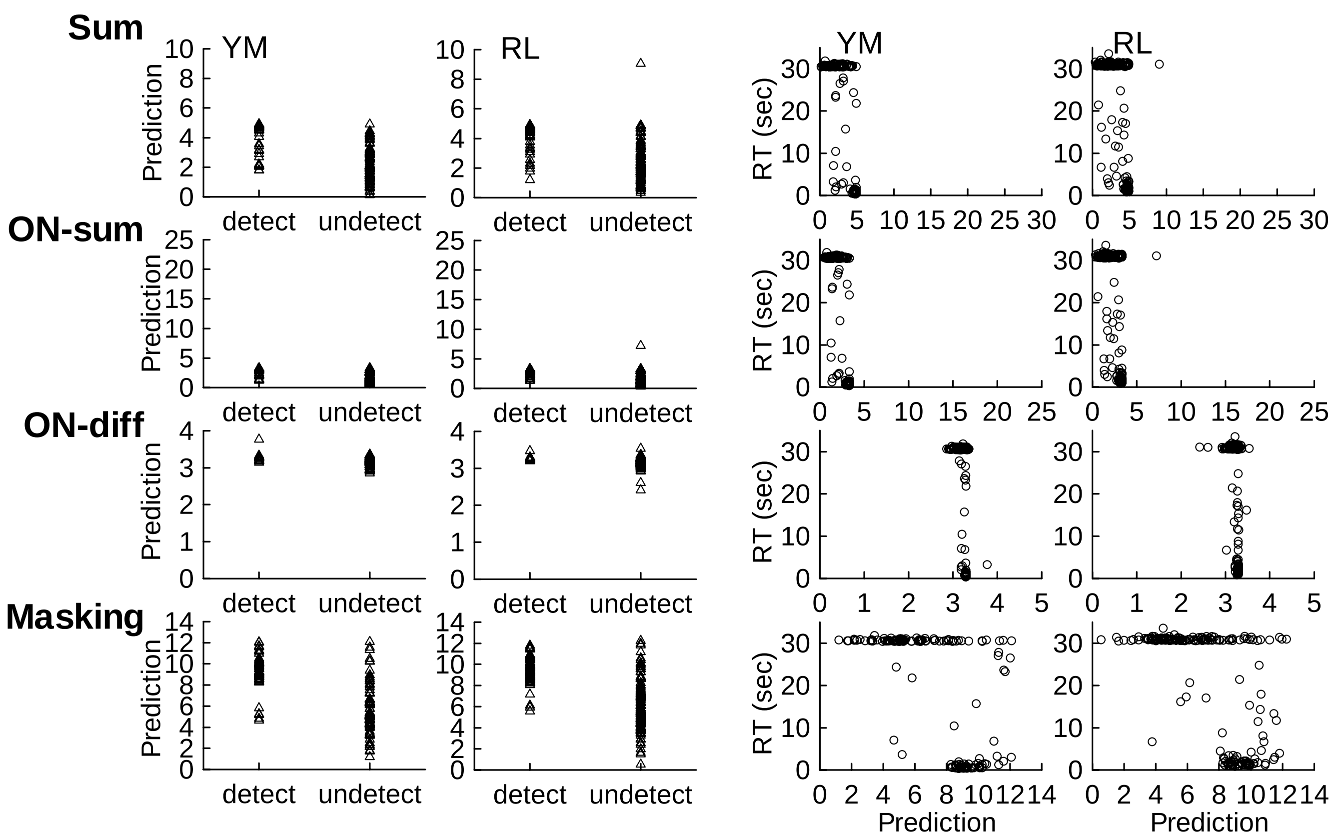


Figure 13. Plots of model predictions vs. detection data for the random position condition.

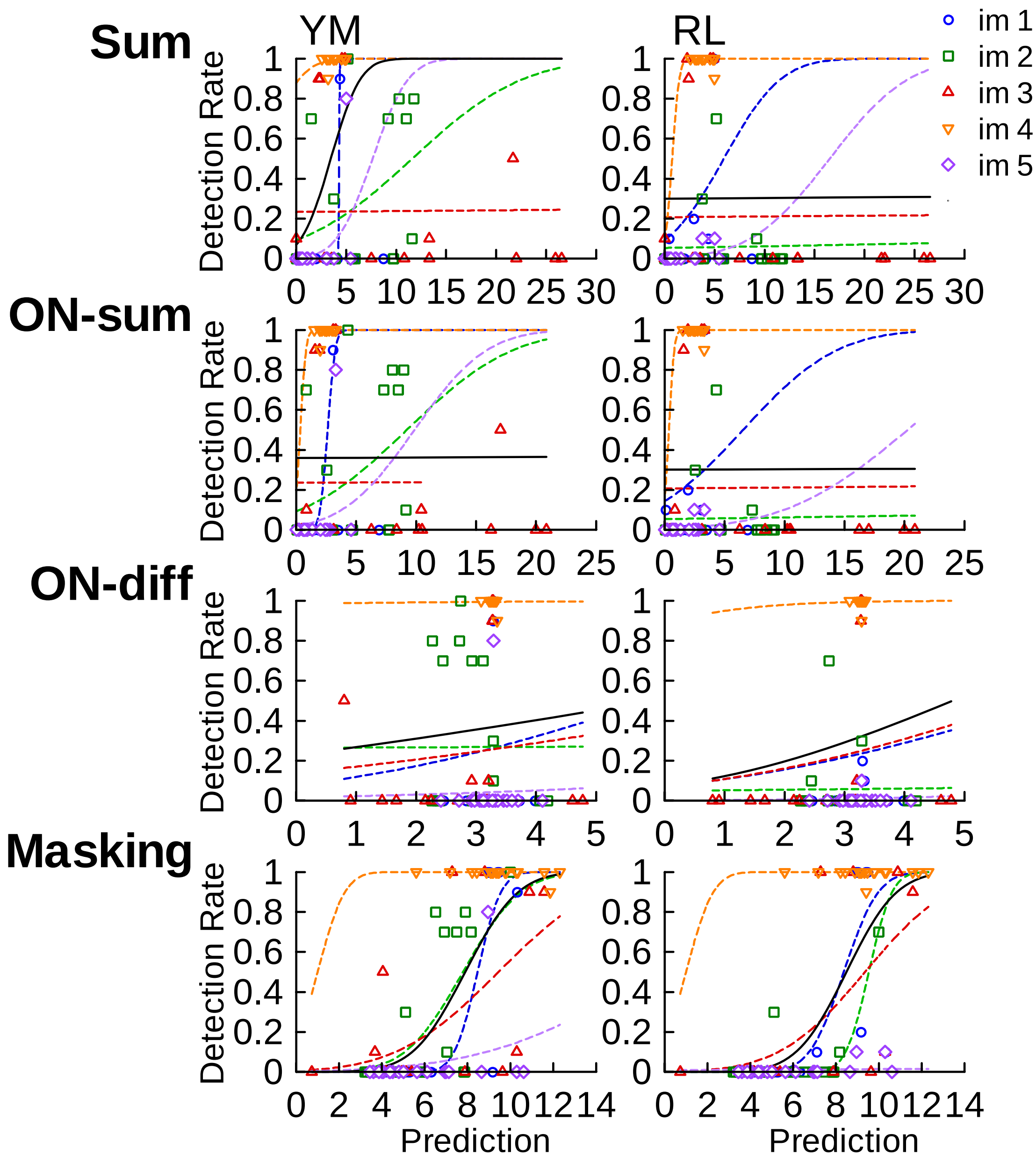


Figure 14. plots of detection model predictions vs. detection performance for two subjects.

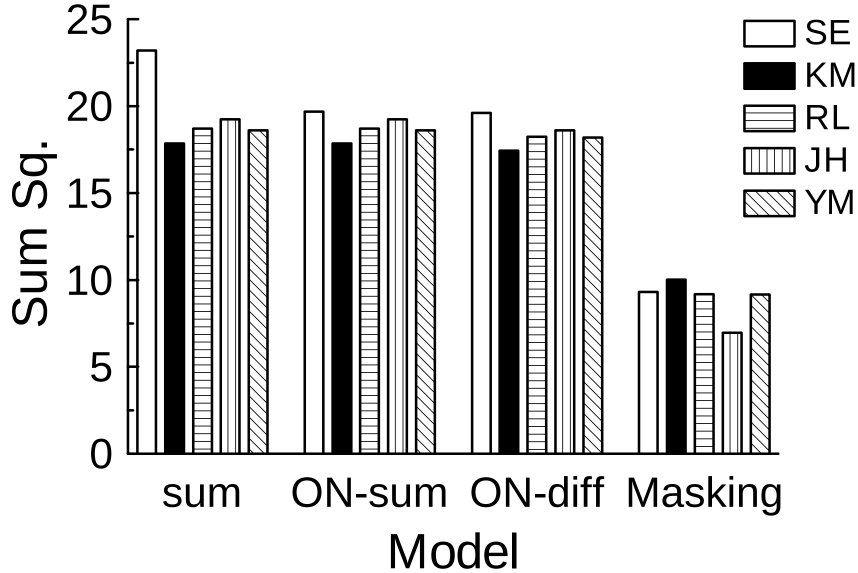


Figure 15. Sum of squared error for fits of four different models of detection with actual detection data for 5 different subjects. The masking model (Ahumada and Beard) appeared to better account for the data.

Lighting effects on detection

Anti collision lights have long been improved for detection purposes and are most useful at night. However their usefulness in terms of improving detection during the day in the context of background masking has not been fully investigated. Part of the utility of using such lights is that when they are abruptly modulated as with strobe lights they create strong stimulation of a highly sensitive visual pathway that is specialized to detect transient changes and motion. This stimulation greatly improves the detection of targets associated with the flashing lights. It has been suggested that apparent motion created by the asynchronous flashing of lights near each other (phi motion) would provide additional advantage for this pathway. The strength of phi motion is determined by the temporal and spatial parameters of the flashing lights. Whether or not phi motion can improve detection beyond that already provided by the transient flashing of the light over the spatial and temporal ranges of typical aircraft has not been tested directly. We performed experiments to measure the effects of flashing and synchrony on detection for simulated aviation targets.

Methods

Figure 16 shows the stimulus conditions including the masking background, the temporal profile of the lights and the target configuration. The subjects' task was choose which quadrant contained the target. In the first experiment we looked at the effects of temporal frequency of the flashing on three modes of lighting: 1) steady wherein both lights stayed on, 2) same, wherein the lights flashed in synchrony, and flutter wherein the light flashed out of phase with each other.

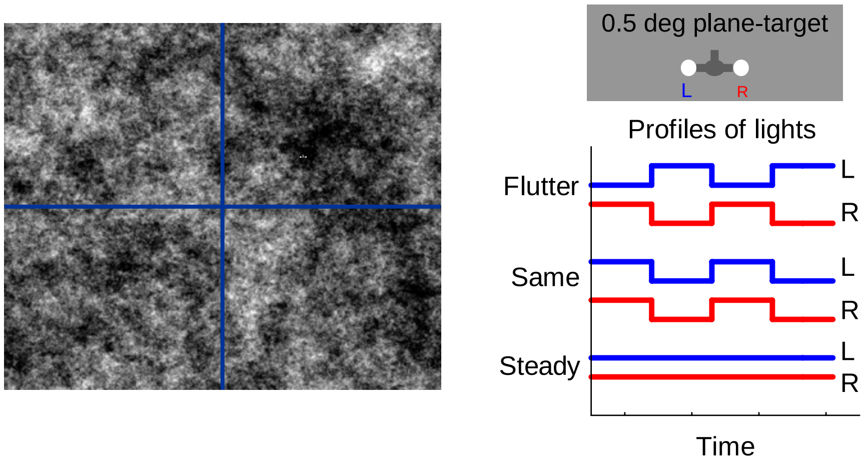


Figure 16. Stimulus conditions for the dual flicker experiments. The masking background is shown on the left, the target ios shown on the upper right and the temporal profile for the three modes of lighting are shown on the bottom right.

The results from the rate experiment are shown in figure 17. All subjects were most sensitive at about 4 Hz where thresholds are minimal. The steady condition shown as the open square was always the least detectable but there was no apparent difference between the flutter and same modes.

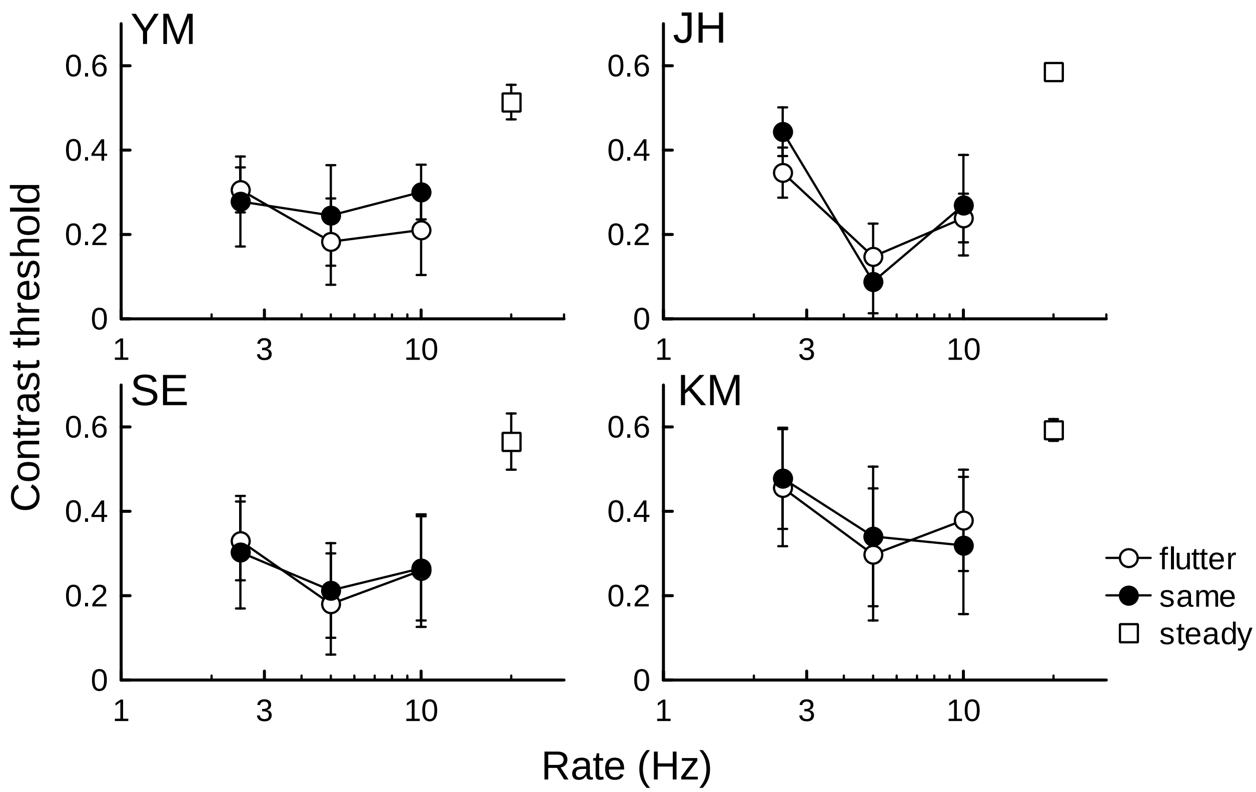


Figure 17 Effects of strobe frequency on detection shown for 4 subjects.

The results from an experiment which looked at the effects of strobe separation distance are shown in Figure 18. Over the range of separations tested, which are reasonable for detection of aircraft, show no appreciable effects other than from mode of lighting wherein again the steady condition was inferior for detection.

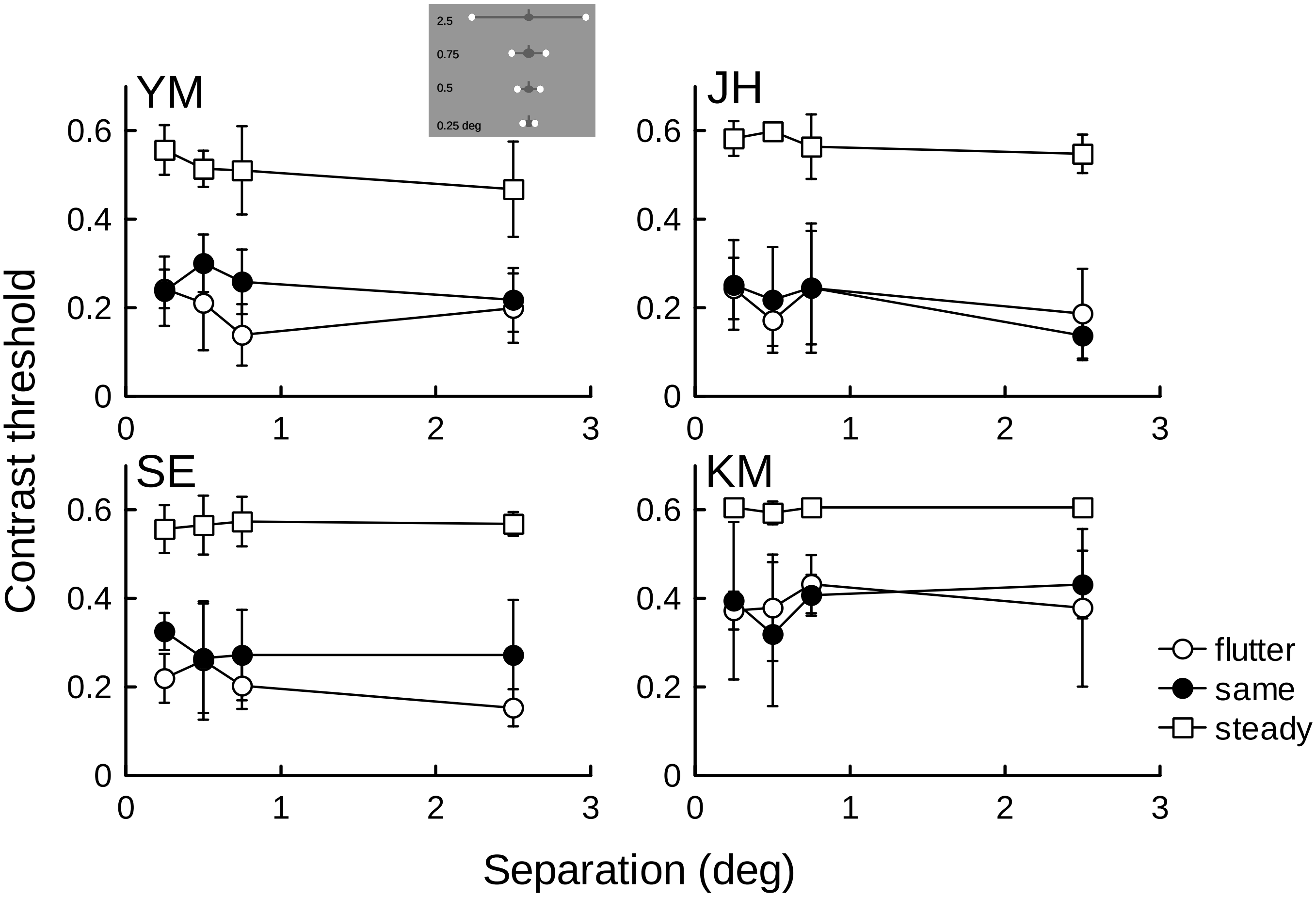


Figure 17. Detection thresholds plotted as a function of strobe light separation in degrees of visual angle.

A third experiment looked at the effects of phase of the flashing lights as it was reasoned that perhaps there was an optimal phase separation for asynchronous lights that produced the most salient phi motion that was not captured in the 180 deg. shift of experiments 1 and 2. The results from this experiment are shown in figure 18. Contrast thresholds were lowest (highest sensitivity) for the 5 Hz stimuli with a 90 degree phase shift appearing optimal although the phase differences did not reach statistical significance.

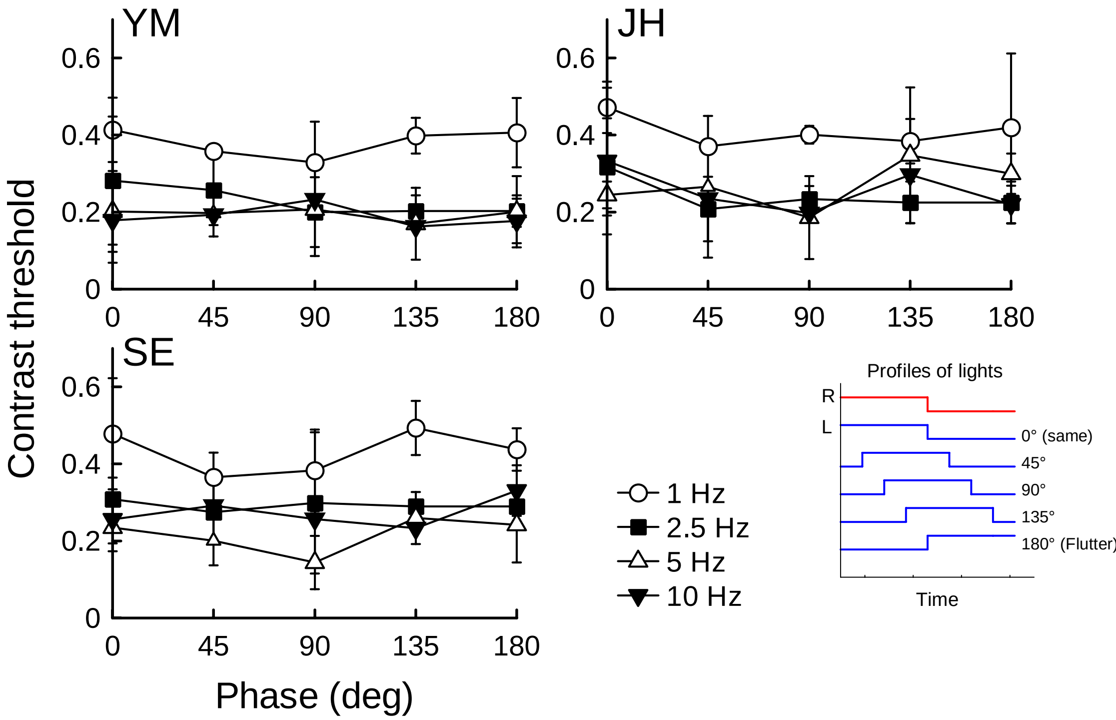


Figure 18. Detection thresholds plotted as a function of phase of offset for the test lights at 4 different flicker frequencies.

Learning to see

As described above optimal visual detection requires prior knowledge of the visual target . We have developed a simple reference card for use in the cockpit (see figure 19). This card illustrates the apparent sizes of typical small airplanes (e.g. Cessna 172) and airliners (e.g. Airbus A-320) at different distances from 2 miles to ½ mile. This card can be used by the pilot to estimate the approximate size of a known but undetected target. Feedback on the use of this card has been quite positive and we will continue to provide it to pilots as requested.

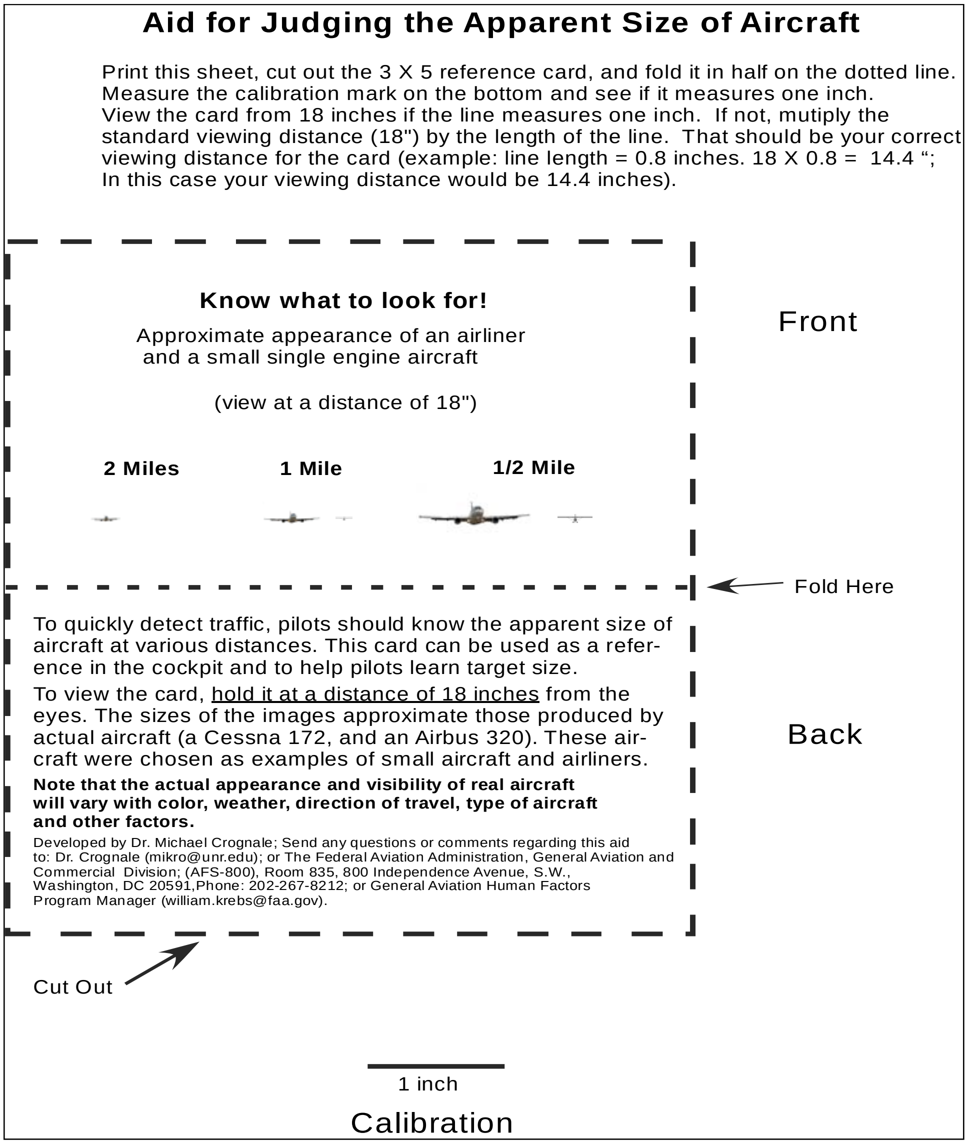


Figure 19.Cockpit aid for aircraft target detection.

We have also developed a proof-of-concept interactive computer program to teach p[ilots how to see better through prior knowledge. The program is based on traffic calls by ATC such as “Cessnna 1234 traffic 2:00 and 2 miles, southbound. Altitude indicates 5,500 ft.” The first part of the program introduces the concept of visibility in the context of the aviation environment. The second part introduces 4 problem areas:1) learning to see; 2) VFR fight into IMC; 3) background masking; and 4) flat light. The third part will be interactive training in two main areas 1) learning to see other aircraft and 2) learning to evaluate the visual environment. The first part will cover judgments of distance, direction, altitude, flight path and orientation. The second part will cover judgments of background masking effects, atmospheric haze , VFR into IMC, and flat light recognition.

We have completed a preliminary version of the part of the program that trains pilots how to judge the appearance and elevation of aircraft traffic given the distance, direction of flight, and altitude from a simulated traffic call. The trainee is also given an altimeter readout and a directional gyro readout in order to provide information to compute relative orientation and altitude. The trainee’s task is to pick the visual scenario that matches the traffic call, out of four possible scenarios that appear on the screen simultaneously. The trainee is also provided feedback to improve learning. This program has initially proven to be useful to improve detection through preliminary results. Future goals include a polished easy to use program that could be distributed to pilots on CD or via download.

Other Products

FAA Handbook Chapter.

We have submitted a manuscript to the FAA describing light, the human visual system, and its limits, and strategies for preflight planning for optimal visual detection, and strategies for seeing and being seen.. Portions of this work were incorporated into the "General Aviation Pilot's guide to preflight weather planning, weather self-briefings, and weather decision making" compiled by Susan Parsons of the FAA. The manuscript will be submitted in its entirety to an aviation Safety publication.

Powerpoint Educational Seminar.

An educational lecture in the form of a Powerpoint presentation has been developed to aid in the educational goals. This presentation includes discussions of light phenomenon, the human visual system and its limits, top-down processing, masking, adaptation, and strategies for optimizing visual performance and for being seen. We have twice presented the educational seminar incorporating our results at Oshkosh. Feedback from The attendees and the National Association for Flight Instructors was quite positive and I have been invited to present again next year. We are also scheduled to present this seminar at Sun n'Fun in Lakeland Florida in April of 2007. A copy of this presentation is included on CD with this report.

References

Ahumada A.J (1996) Simplified Vision Models for Image Quality Assessment. SID International Digest of Technical Papers, Volume XXVII, May, 397-400.

Ahumada A.J. & Beard B.L. (1997) Image Discrimination Models Predict Detection in Fixed but not Random Noise. JOSA A, 14, 2471-2476.

Ahumada A.J. & Beard B.L. (1998) A simple Vision Model for Inhomogeneous Image Quality Assessment. SID Digest of Technical Papers XXIX, paper 40.1.

Atick, J.J. & Redlich, A. N. (1992). What Does the Retina Know about Natural Scenes? Neural Computation, 4, 196-210.

Field D.J. & Brady N. (1997) Visual Sensitivity, Blur and the sources of Variability in the Amplitude Spectra of Natural Scene. Vision Research 37, 3367-3383.

Khatwa R. & Roelen A.L.C. (1996) An Analysis of Controlled-flight-into-terrain (CFIT) Accidents of Commercial Operators 1988 through 1994, Flight Safety Foundation, Pilot Safety Digest, Special Issue, April-May.

Krause R.R (1995) Avoiding Mid-Air Collisions. S.S. Krause, Tab Books.

Mizokami Y. & Crognale M.A. (2005a) Detection of dual flashing lights. *Annual meeting for the Vision Sciences Society*, Sarasota, FL.

Mizokami Y. & Crognale M.A. (2005b) Detection model predictions for aircraft targets on natural backgrounds. European Conference on Visual Perception.

O'Hare, D. & Owen D (2002) Cross-country VFR crashes: pilot and contextual factors.. Aviation Space and Environmental Medicine (73) 363-366.

Olshausen, B.A., & Field D.J. (1996). Emergence of Simple-Cell Receptive Field Properties by Learning a Sparse Code for Natural Images. Nature, 381, 607-609.

Olshausen, B.A., & Field D.J. (1997). Sparse Coding with an Overcomplete Basis Set: A Strategy Employed by V1? Vision Research, 37, 3311-3325.

Rohaly A.M., Ahumada, A.J., Watson A.B. (1997) Object Detection in Natural Backgrounds predicted by Discrimination Performance and Models. Vision Research 37, 3225-3235.

Simoncelli, E.P & Olshausen B.A. (2001) Neural Image Statistics and Neural Representation, E.P. Annual Reviews of Neuroscience, 24,1193-1216.

Volpe (1994) Controlled Flight Into Terrain (CFIT) Report, 1983-1994